Live Migration of Guest Computers Using the NEMO BS Mobility Function

In order to construct an efficient virtual computing environment, a system for ensuring the availability of virtual computers and managing them flexibly is required. In this whitepaper a virtual computer migration method using NEMO BS technology, which adds a mobility function to IPv6 routers, is discussed and the results of experiments using this setup are examined. NEMO BS makes it possible to migrate virtual computers beyond the network segment.

3.1 Background

The capability of individual computers is progressing rapidly as a result of the processing power of computers and technological advances in networks and data stores. Meanwhile, there is also the view that speed improvements for individual computers will slow down in the future. In recent years cloud technology that treats multiple computers as a single computer resource is gaining a lot of attention*1*2 as a remedy for this issue. On the other hand, the virtual computing technology that allows a single computer resource to be partitioned virtually and used as multiple different computers*3 has been studied for many years. These technologies may appear at a glance to be targeted at different uses, but it is believed that they can supplement each other and make it possible to utilize computer resources more efficiently. For example, Amazon's EC2*4 service provides a cloud environment that combines multiple computer resources into a single service, but virtual computer technology is also used in each individual computer resource. Using units broken down into small virtual computers instead of using large resource units consisting of individual physical computers enables the efficient and flexible use of computer resources, and consequently provides convenience in excess of the overhead created by partitioning into virtual computers.

A system for managing virtual computers flexibly is crucial for developing this kind of virtual computing environment. This is due to the fact that the ability to arrange the necessary amount of virtual computers in the necessary location in accordance with the requirements of a cloud service contributes to overall performance increases and the efficient use of resources. In this whitepaper, systems for relocating virtual computers are focused on, and a technique for migrating running virtual computers to a different segment (offlink segment) is proposed.

3.2 Live Migration Issues

Currently, a large number of virtual computer technologies are provided at a practical level. Some of these provide functions for migrating virtual computers from their parent computer (*host computer*) to another host computer. VMware's VMotion and Xen's*5 Live Migration are examples of this kind of technology. In this whitepaper, virtual computers that are generated in the host computer are called *guest computers*, and the function for migrating running guest computers between host computers is called *live migration*. The use of a live migration function makes it possible to migrate running guest computers to another machine with very little downtime. However, the functions that are currently provided are restricted to situations in which the source host computer and destination host computer belong to the same segment.

This restriction is due to the network configuration method provided to guest computers. Host computers and guest computers do not have an equal relationship. Host computers normally have control over all resources, and allocate part of these resources to guest computers. When connecting a guest computer to a network, a configuration like the one shown in Figure 1 is used.

*1 Aaron Weiss. Computing in the clouds. netWorker, Vol. 11, No. 4, pp. 16-25, December 2007.

*2 Brian Hayes. Cloud computing. *Communications of the ACM*, Vol. 51, No. 7, pp. 9-11, July 2008.

*3 Paul Barham, Boris Dragovic, Keir Fraser, Steven Hand, et. al. Xen and the Art of Virtualization. In *SOSP '03: Proceedings of the Nineteenth ACM Symposium on Operating Systems Principles*, pp. 164-177, ACM, 2003.

*4 Amazon. Amazon Elastic Compute Cloud (Amazon EC2), October 2009. http://aws.amazon.com/ec2/

*5 Citrix, October 2009. http://www.xen.org/

Using configuration (a) in Figure 1, the guest computers are connected to the same network as the host computer via a virtual switch provided by the host computer. Similarly, a virtual switch is also created between the host computer and the guest computers when using configuration (b). Unlike configuration (a), under this configuration the host computer also functions as an upstream router for the guest computers. As Figure 1 clearly shows, the network configuration for the guest computers is highly dependent on the host computer. When executing live migration, no changes are made to the operating environment of the guest computer itself. Consequently, it is only possible to carry out live migration using configuration (a). Under configuration (b), the address allocated to the virtual switch differs depending on the location where the host computer is attached. This means that after a guest computer is migrated, communications cannot be continued until the network environment of the guest computer is updated appropriately. The same problem also occurs under configuration (a) if the source and destination host computers for migration are connected to different segments.

Live migration makes it possible to avoid consolidating virtual computers on a single host computer, but this function is limited to migration within the same segment. Even if there is a host computer on another segment with free resources, it cannot be utilized. It is also not possible to relocate resources to improve performance, such as migrating a guest computer to a host computer that is closer to the user.

3.3 Overview of NEMO BS

NEMO BS (Network Mobility Basic Support)*6 is a protocol for adding a mobility function to IPv6 routers. In a NEMO BS environment, a *MR* (*mobile router*) compatible with NEMO BS manages the *MNP* (*mobile network prefixes*) that serve as fixed network prefixes. IPv6 hosts connected to the network provided by the MR use fixed addresses in the MNP range. The MR connects to various segments over the Internet, and maintains connectivity with the Internet according to the target network environment. During this the MNP the MR manages does not change. Hosts attached to the network controlled by the MR are always able to maintain the same network environment regardless of the location of the MR.

This function is achieved with the support of the *HA* (*home agent*) that serves as a counterpart to the MR (Figure 2). The MR acquires an address (care-of address) on the destination network depending on the network environment, and establishes a bidirectional IPv6 over IPv6 tunnel with the HA. Traffic originating from nodes within the MNP is sent to the HA using this tunnel, then transmitted from the HA to the target of communications. Meanwhile, traffic directed to nodes within the MNP is received by the HA, delivered to the MR via the tunnel, then transmitted to the final destination host.

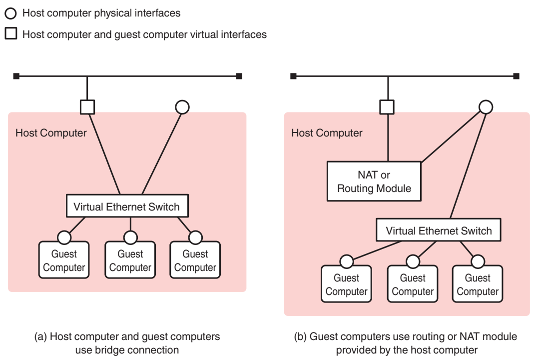

Figure 1: Guest Computer Network Configurations

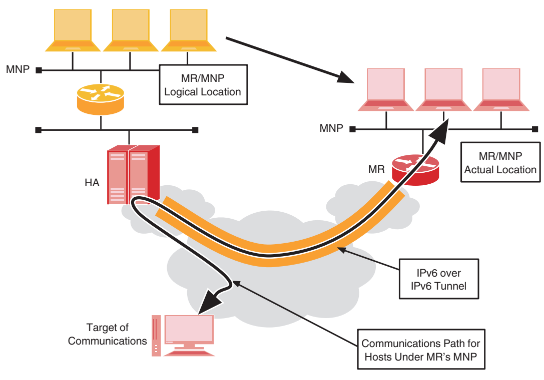

Figure 2: Operational Overview of NEMO BS

*6 Vijay Devarapalli, Ryuji Wakikawa, Alexandru Petrescu, and Pascal Thubert. Network Mobility (NEMO) Basic Support Protocol. IETF, January 2005. RFC3963

3.4 Design

Live migration is limited to migration within the same segment because the network environment for guest computers is dependent on the host computer. In other words, if the guest computer's network environment is designed so that it is fixed and does not depend on the host computer, it will be possible to migrate guest computers to host computers connected to different segments.

In this whitepaper, a method for using IP mobility technology to maintain a fixed network environment on guest computers is proposed. There are two approaches to this: equipping the guest computers with a host mobility function such as Mobile IP[*7*8], and utilizing a system such as NEMO BS on the host computer to provide a fixed network for guest computers. The former option requires the modification to guest computers (the installation of an IP mobility function), but this makes the migration of individual guest computers possible, allowing for finer control over computer resources. The latter option allows existing guest computers to be used without any modifications, but requires that the migration of guest computers and the migration of the host computer serving as the MR for NEMO BS be synchronized. In this whitepaper the latter option is focused on, assuming a scenario where the guest computers used in an existing system will continue to be utilized.

The system design may differ depending on the resource management method for host computers. In this whitepaper a design using the virtual computing environment provided by Xen is discussed, but it should be possible to apply a similar design to other systems without significant changes. Figure 3 shows an overview of this design. The configuration used is an expanded version of (b) from Figure 1. The host computer functions as the MR, providing two interfaces, one is for connecting to the Internet and the other is for providing MNP. The MNP connection interface is a virtual interface rather than a physical interface. This virtual interface is connected to the virtual switch that the host computer provides for guest computers. The host computer's virtual interface and the interfaces for guest computers connected to the virtual switch are allocated fixed addresses managed by the MR. Through the NEMO BS function, addresses within MNP do not change regardless of where the host computer is physically connected.

In the proposed environment for live migration of guest computers, multiple host computers are located on the network. These host computers act as MR and are configured with the same MNP, but host computers that have no active guest computers do not function as an MR. When migrating guest computers, regular live migration procedures are first performed to migrate guest computers to the destination host computer. At this point, guest computers are still disconnected from the network. Following this, the source host computer sends a migration complete notification to the destination host computer, and suspends its NEMO BS function. The host computer that receives the migration complete notification registers its location with the HA, and begins operation as a NEMO BS MR. Once registration with the HA is complete, the MNP for the virtual switch that the guest computers are connected to is activated, and migration is complete.

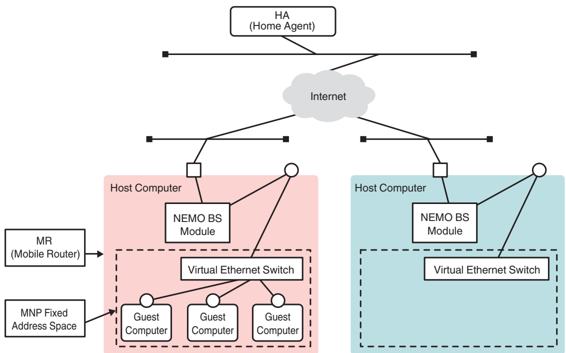

Figure 3: Guest Computer Migration Using NEMO BS

*7 Basavaraj Patil, Phil Roberts, and Charles E. Perkins. IP Mobility Support for IPv4. IETF, August 2002. RFC3344

*8 David B. Johnson, Charles E. Perkins, and Jari Arkko. Mobility Support in IPv6. IETF, June 2004. RFC3775.

3.5 Verification Tests

To confirm that the proposed design is feasible, a prototype system was implemented and operation experiments were performed. The experiment environment was configured with a total of five computers: two Xen host computers operating as MR, a computer with HA function, a computer for receiving streaming data during testing, and a control computer for initiating migration operations. A guest computer was created with the Xen host computers and configured to operate as a streaming server.

In order to carry out experiments in an environment close to real conditions, this equipment was placed into an actual Internet environment. The fixed network for the HA and MR to use was installed as a part of the network constructed for Interop Tokyo 2009*9, and the Xen host computers and stream data reception nodes were placed on the IIJ network. Each piece of equipment was connected to the others via the Internet. Figure 4 shows an overview of the experiment network.

15 MB, 520 KB/s MPEG4 stream data was continually sent from the guest computer (the streaming server) using UDP. The instructions to begin live migration of the guest computer, and the suspension and initiation of the host computer NEMO BS function were carried out over SSH using a separate control computer. The guest computer live migration command was executed remotely from the control computer, and host computer NEMO BS movement processing was executed immediately after live migration was complete. The guest computer was repeatedly migrated between the two host computers every five minutes.

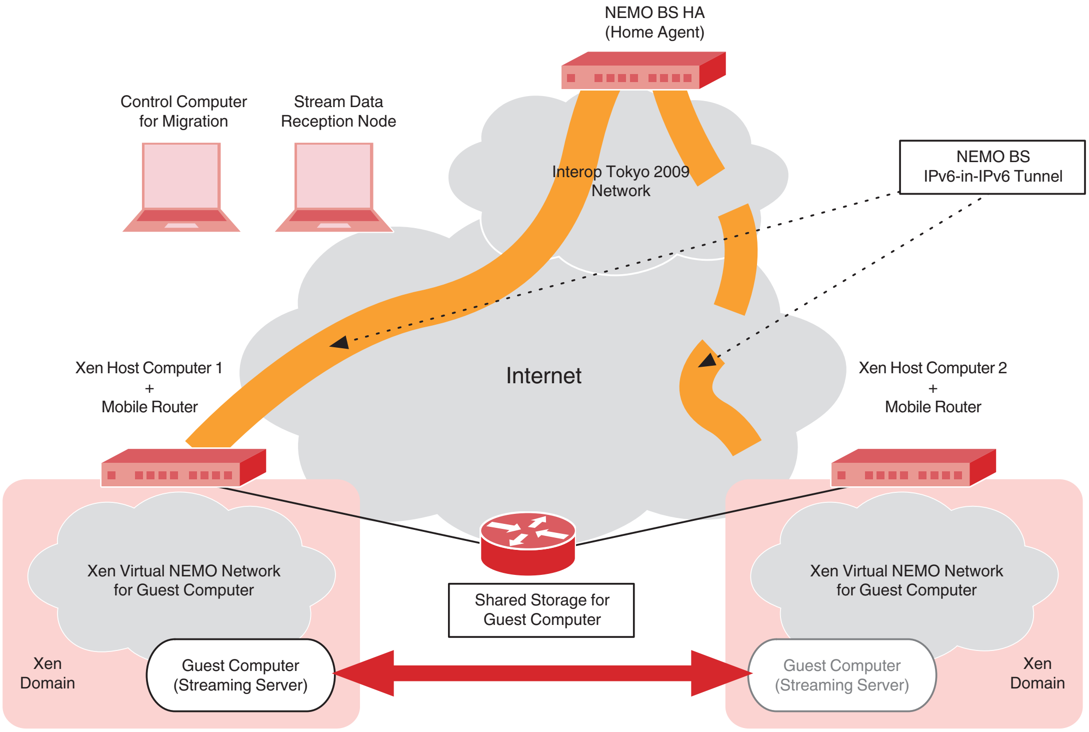

Figure 4: Experiment Network Overview

*9 Interop Tokyo 2009, June 2009. http://www.interop.jp/

3.6 Evaluation and Issues

Figure 5 shows the results of monitored stream traffic over the host computer's Internet interface. Live migration of the guest computer was carried out every five minutes, and migration between the two host computers can be seen as a movement of the traffic between two host computers.

When using live migration, the source virtual computer continues operation while virtual computer data is being copied to the destination host computer. The downtime when operation is switched over is quite small (about several hundred milliseconds). However, as network layer mobility technology was used to migrate the host computer in this proposed mechanism, node movement processing was also necessary for the host computer after the guest computer was migrated. When examining the traffic data details, it became apparent that it was taking approximately six seconds for the destination host computer to begin sending streaming data after the source host computer had suspended the distribution of streaming data.

As router advertisement messages on the network where the host computers connect were made at three to four second intervals, the host computers were in an environment where they could receive router advertisement messages in about two seconds. Even when the one second care-of address duplicate confirmation that is a part of NEMO BS processing is taken into account, movement processing should be completed in approximately three seconds in this environment. The following two points are possible reasons for processing taking longer than this in the experiment.

1. Overhead due to the execution of the remote control script from the control computer.
2. Procedures that differ from regular movement processing procedures.

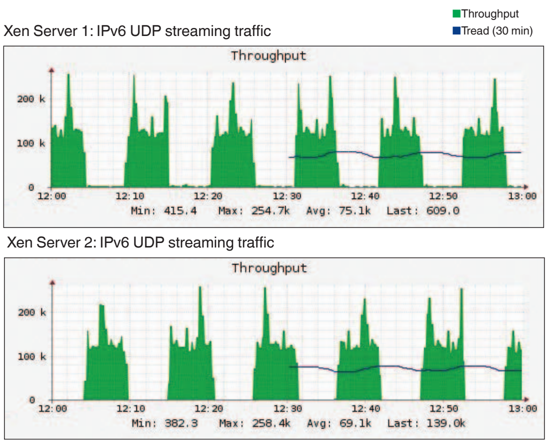

Figure 5: Stream Traffic Transition

As the suspension and initiation of the NEMO BS function were executed using a remote script via SSH, more time than usual was required for actions such as TCP connection establishment, process generation on the host computer, and controlling the NEMO BS daemon program via remote command. Additionally, movement processing that would normally be executed on a single mobile node was jointly executed on two different mobile nodes in the proposed mechanism. In other words, one MR executed registration processing after the other MR finished cancelling registration with the HA. When carrying out movement processing using the same MR, there is no need to cancel registration of old information when registering new information, so updates are made consecutively. In the proposed mechanism extra time was required for canceling registration.

3.7 Consideration

The prototype implementation and experiments aimed at verifying the combination of two technologies (NEMO BS and Xen virtualization). This section mentions some observations found during the experiments.

3.7.1 Network Storage Issues

There are two methods of providing storage for guest computers under systems that offer a virtual computing environment, one is providing part of the host computer's storage, and the other is providing network storage. Storage must be located on the network when using live migration, as the host computer that accommodates guest computers will change. As guest computers are not migrated beyond the same segment under current assumed usage patterns, there are no significant changes to the environment before or after migration even when network storage is used. However, when migrating guest computers beyond their segment as in the proposed mechanism, there is the possibility that network storage reachability and communication delay issues may occur.

One solution to this is constructing a guest computer environment that carries out all processing in memory without using external storage. However, under this method increases in the amount of data to be stored in memory will cause migration to take longer to complete. This construction also makes it difficult to handle large amounts of data.

Another method is to provide storage that can be handled efficiently and transparently from multiple locations on the Internet. For example, developing storage mirroring technology and making it possible to access the same storage data from multiple locations should ease the network quality changes to guest computers and also allow for improvements to be made to the fault-tolerance of storage.

3.7.2 Usability Issues

In the proposed mechanism the NEMO BS technology is used, but there are two issues with this method.

First, there is the fact that tunnel communications via HA is a prerequisite, although the same is true for all Mobile IP-based mobility technologies in general, and not just NEMO BS. This creates new issues to be considered, such as the HA becoming a single point of failure, and the need to introduce redundancy technology into the HA. It may be possible to resolve these issues by adopting different mobility technology that does not use tunneling, such as HIP[*10], LIN6[*11], or MAT[*12]. To achieve the offlink live migration of virtual computers, the primary requirement is to maintain the network environment of the virtual switch that virtual computers connect to. Consequently, the method used to achieve this does not need to be NEMO BS.

The second issue is that by using NEMO BS, all guest computers are managed as a cluster of nodes attached to the same fixed network. This means that when a host computer is migrated, all related guest computers must also be migrated at the same time. This issue can be dealt with using one of the following methods. One method is to have guest computers adopt a host-based mobility technology, for example Mobile IP. In this case all guest computers must support Mobile IP, increasing the requirements for deployment. NEMO BS has the large advantage of it not being necessary to modify guest computers, so standard guest computers can be used as-is. Providing a NEMO BS environment implemented as a virtual computer for each guest computer would be a compromise for this. This would operate as a kind of Mobile IP proxy. This method would make it possible to migrate individual guest computers while avoiding the need to make any changes to them.

3.8 Conclusion

Ensuring the mobility of computer resources will be crucial in the cloud computing environments of the future. Virtual computer live migration technology is a strong candidate for ensuring resource mobility. However, using current live migration technology the migration destination is restricted to other host computers within the same segment. This is because the network environment provided to guest computers is dependent on the physical network that host computers connect to. If this restriction could be removed and migration to other remote segments were made possible, it would allow for more flexible management of computer resources. In this whitepaper a technique for migrating guest computers between multiple host computers that do not share a segment using virtualization and NEMO BS technologies is proposed. The use of NEMO BS technology to provide guest computers with a fixed network environment regardless of the network a host computer connects to makes cross-segment migration possible.

In closing, while conducting this research I received a great deal of advice regarding Linux and MIPL/NEPL configurations from Masahide Nakamura, Jean Lorchat, and Martin André. I would also like to thank Takashi Miyake, Manabu Ori, and the Interop Tokyo 2009 NOC team for their assistance in preparing the test environment.

Author:
Keiichi Shima
IIJ Innovation Institute Inc. Research Laboratory
Mr. Shima is pursuing the research and development of IP mobility technology that will be necessary for transitioning continually evolving Internet terminals to wireless systems.

*10 Robert Moskowitz, Pekka Nikander, Petri Jokela, and Thomas R. Henderson. Host Identity Protocol. IETF, April 2008. RFC5201

*11 Mitsunobu Kunishi, Masahiro Ishiyama, Keisuke Uehara, Hiroshi Esaki, and Fumio Teraoka. LIN6: A New Approach to Mobility Support in IPv6. In Wireless Personal Multimedia Communication (WPMC), November 2000.

*12 Reiji Aibara, Takahiro Fujita, Kaori Maeda, and Yoshihiro Nomura. Mobile Internet Architecture with Address Translation. Special Issue on Next Generation Mobile Communication Networks and their Applications. IPSJ Journal Vol. 43, No. 12, pp. 3889-3897, December 2002.